

П. П. ЕРМОЛОВ, канд. техн. наук, Севастопольский национальный технический университет

ФОРМАЛИЗАЦИЯ СТРУКТУРЫ ПОРТАЛА ЗНАНИЙ ПО ИСТОРИИ НАУКИ И ТЕХНИКИ

Проведено формализованное описание структуры портала знаний по истории науки и техники. Построена базовая онтология портала, в которую входят 11 классов: «исследователи», «методы исследования», «источники», «события», «организации», «географическое место», «мемориальный объект», «период», «научный результат», «раздел науки» и «объект исследования». Последние три класса представлены как метапонятия предметной онтологии портала. На классах и подклассах определены отношения наследования, включения, отношения «класс — данные», выделены 30 наиболее актуальных ассоциативных отношений. Для упрощения создания онтологии и реализации релевантного поиска предложено деление крупных источников (монографий и обзоров) на более мелкие фрагменты и организация подкласса «цитаты». При проведении формализации использованы аппарат теории множеств и «двухслойная» концепция историографии науки и техники как метанауки.

Ключевые слова: знание, история науки и техники, базы знаний, методы исследования, базы данных

В настоящее время значительно активизируются исследования, которые находятся на стыке с Интернетом. В научную терминологию прочно вошли такие понятия, как «семантический Веб», «интеллектуальные сети», расширяется теоретико-методологический аппарат знаниеориентированных наук и технологий. В данном контексте следует отметить направления исследований онтологий, которые (к этому мнению склоняется большинство специалистов) составят основу развития следующих поколений Интернета, включая идею семантического Веба. Различают до девяти задач, которые решаются с использованием онтологий [1]:

1. Создание и использование баз знаний;
2. Организация эффективного поиска в базах данных, информационных каталогах, базах знаний;
3. Создание систем, реализующих механизмы рассуждений;
4. Организация поиска по смыслу в текстовой информации;
5. Семантический поиск в Internet;
6. Представление смысла в метаданных;
7. Построение и использование баз общих знаний для различных интеллектуальных систем;
8. Обеспечение общей терминологии для множества специалистов и совместно используемых приложений;
9. Многократное применение баз знаний и информационных массивов, представляющих сведения о технических системах на различных стадиях их жизненного цикла.

К перечисленным задачам можно добавить следующую:

10. Многократное применение информационных массивов, представляющих сведения о результатах научных исследований по отдельному узкому направлению исследований.

В отличие от общей постановки [2], деление на базовые и предметные онтологии в задачах, касающихся научных исследований по отдельному узкому направлению, производится с учетом специфики предмета исследования. Если ограничиться такой областью, как историография науки и техники, то кроме традиционно относимых к числу базовых элементов онтологии классов ИССЛЕДОВАТЕЛИ (персоны), ИСТОЧНИКИ (публикации), СОБЫТИЯ, ОРГАНИЗАЦИИ, ГЕОГРАФИЧЕСКОЕ МЕСТО, к этому числу следует отнести также класс МЕТОДЫ ИССЛЕДОВАНИЯ: нами принята классификация методов исследования, предложенная проф. Смоленским Н. И., в соответствии с которой выделены описательно-повествовательный, биографический методы и метод терминологического анализа [3, 4]. В последнее время в историографии науки и техники получает распространение также библиометрический подход [5], в связи с чем три перечисленных выше метода дополнены методом библиометрического анализа.

Спецификой такого ограничения портала знаний диктуется и деление крупных источников (экземпляров класса ИСТОЧНИКИ, к разряду которых можно отнести монографии и обзоры по тематике портала), на более мелкие фрагменты (подкласс Цитаты). В результате такого деления, во-первых, упрощается создание онтологии, и, во-вторых, поиск по portalу становится более релевантным. Кроме этого, специфичными для базовой онтологии портала в области историографии науки и техники являются понятия (классы) МЕМОРИАЛЬНЫЙ ОБЪЕКТ и ПЕРИОД. К базовой онтологии относятся также традиционные для этой категории метапонятия предметной онтологии: НАУЧНЫЙ РЕЗУЛЬТАТ, РАЗДЕЛ НАУКИ и ОБЪЕКТ ИССЛЕДОВАНИЯ, которые выделяют в рассматриваемой области знаний (историографии развития конкретной области науки или техники) наиболее значимые разделы и подразделы, задают типизацию объектов исследования и описывают результаты научной деятельности. Эти понятия (классы) упорядочиваются в иерархию общее — частное или часть — целое.

Цель настоящей работы — построение базовой онтологии портала знаний по истории науки и техники. Для достижения этой цели привлечен аппарат теории множеств и использована «двухслойная» концепция историографии науки и техники как метанауки [6].

Формально онтология портала знаний представляет собой кортеж вида

$$O = \langle C, R, T, D, A, F \rangle,$$

где $C = \{C_1, \dots, C_n\}$ – конечное непустое множество классов, описывающих понятия некоторой предметной или проблемной области; $R = \{R_1, \dots, R_m\}$,

$R_i \subseteq C \times C$, $R = \{R_T\} \cup \dots \cup \dots$ – конечное множество бинарных отношений, заданных на классах (понятиях): R_T – отношение наследования, R_p – отношение включения («часть – целое»), R_A – конечное множество ассоциативных отношений; T – множество стандартных типов данных; $D = \{d_1, \dots, d_n\}$, – множество доменов $d_i = \{s_1, \dots, s_k\}$, где s_i – значение стандартного типа из T ; $TD = T \cup \dots$ – обобщенный тип данных, включающий множество стандартных типов и множество доменов; $A = \{a_1, \dots, a_w\}$, $A \subseteq C \times TD \cup \dots$ – конечное множество атрибутов, описывающих свойства понятий C и отношений R_A, F – множество ограничений на значения атрибутов понятий и отношений, т. е. предикатов вида $p_i(e_{i1}, e_{i2})$, где e_{ik} – это либо имя атрибута ($e_{ik} \in A$), либо константа ($e_{ik} \in td_j$, где $td_j \in TD$).

Особенностью отношения R_T является то, что при наследовании от родительского класса его классу-потомку передаются не только все атрибуты, но и отношения. Отношение включения («часть–целое») R_p наделено свойством транзитивности, благодаря этому при поиске объектов можно осуществлять транзитивное замыкание по этому отношению. Набор ассоциативных отношений R_A определяется разработчиком онтологии. Наличие таких отношений позволяет организовать содержательный поиск и навигацию по контенту портала знаний. Важной особенностью отношений R_A является то, что они могут иметь собственные атрибуты, специфицирующие связь между аргументами [2, 7].

Особенностью рассматриваемой нами задачи является построение онтологии в более узкой области знаний. По этой причине в онтологию вводится еще один вид отношений – отношения типа «класс – данные». В итоге онтология задачи может быть представлена следующим образом

$$O = \langle C, R, R_{CD}, T, D, A, F \rangle,$$

где R_{CD} — отношения «класс — данные».

Ранее отмечалось, что в отличие от идеологии создания портала знаний вообще специализированный портал знаний вполне естественно должен отличаться характером и соотношением базовых и предметных онтологий. С учетом этого структуру классов портала знаний по истории науки и техники целесообразно представить в формате классов трех уровней: собственно класса и двух подклассов: 1-го и 2-го уровня. В соответствии с этим, а также с учетом «двухслойной» концепции историографии науки и техники как метанауки [6] формальное выражения

для собственно класса и подклассов 1-го и 2-го уровня соответственно будут выглядеть следующим образом:

$$C^i, C^{iH}, \\ C^{i.j}, C^{iH.j}, \\ C^{i.j.k}, C^{iH.j.k}$$

(где индекс H соответствует тому, что класс или подкласс описывает только историографические аспекты портала), а n -е экземпляры соответствующих классов и подклассов будут иметь вид:

$$C_n^i, C_n^{iH}, \\ C_n^{i.j}, C_n^{iH.j}, \\ C_n^{i.j.k}, C_n^{iH.j.k}.$$

Опишем классы и подклассы базовой онтологии портала знаний по истории науки и техники более подробно (см. рис. 1).

Класс C^1 описывает множество ИССЛЕДОВАТЕЛИ и содержит следующие атрибуты и домены (здесь и далее домены указаны в скобках):

Фамилия,

Имя,

Отчество,

Инициалы,

Страна,

Пол (женский / мужской),

Ученое звание (без ученого звания / старший научный сотрудник / доцент / профессор / член-корреспондент / академик),

Ученая степень,

Награды,

Почетные звания,

E-mail,

Раб. тел.,

Моб. тел.,

Почтовый адрес,

Дата рождения (год, месяц, день),

Дата смерти (год, месяц, день),

Фамилия и инициалы на других языках,

URL.

Подклассы $C^{1.1}$ и $C^{1.2}$ описывают множества исследователей в области техники и исследователей в области историографии соответственно.

Класс C^{2H} описывает множество МЕТОДЫ ИССЛЕДОВАНИЯ и, как отмечалось ранее, принадлежит к элементу базовой онтологии портала знаний по историографии науки и техники. Класс и его подклассы не

содержат экземпляров, описание класса содержит только перечисление методов, описанных в дочерних подклассах:

$C^{2H.1}$ – описательно-повествовательный,

$C^{2H.2}$ – биографический,

$C^{2H.3}$ – терминологического анализа,

$C^{2H.4}$ – библиометрического анализа.

Атрибуты класса C^{2H} и его подклассов:

Другие названия метода,

Описание,

URL,

Дата возникновения.

Класс C^3 описывает множество ИСТОЧНИКИ. Класс содержит следующие атрибуты и домены:

Библиографическая ссылка на источник,

Язык источника (рус. / укр. / англ. / нем. / фр. / исп. / другой),

Тип источника (автореферат / архивный документ / биобиблиографический указатель / директивный материал / диссертация / доклад / инструкция / каталог или буклет / картографическое издание / монография или глава / отчет / патент или а. с. / препринт / реферат / рецензия / рукопись / сборник / справочник / стандарт / статья / тезисы / учебно-методическое издание / е-ресурс),

Дата происхождения,

Дата публикации,

Краткая аннотация,

URL.

Единственным подклассом класса C^3 является $C^{3.1}$ — Цитаты, введенный по упоминаемым выше причинам. Этот подкласс содержит два атрибута:

Ссылка на родительский класс,

Библиографическая ссылка на источник с указанием номера страницы.

Класс C^4 описывает множество СОБЫТИЯ. Класс содержит следующие подклассы:

$C^{4.1}$ – события в области науки и техники,

$C^{4.2}$ – события в области историографии.

Класс $C^{4.1}$ и его подклассы имеют следующие атрибуты и домены:

Название события,

Дата начала события (год / месяц / день),

Дата окончания события (год / месяц / день),

Краткое описание события.

Класс C^5 описывает множество ОРГАНИЗАЦИИ. Класс не содержит подклассов и имеет следующие атрибуты и домены:

Название организации,

Тип организации (ассоциация / библиотека / военный или стратегический объект / издательство / институт АН / компания / музей / неформальная организация / отраслевой НИИ / промышленное предприятие / университет)

Аббревиатура,

Описание организации,

Адрес,

Телефон,

Факс,

e-mail,

Дата основания,

Дата ликвидации.

Класс C^6 описывает множество ГЕОГРАФИЧЕСКОЕ МЕСТО. Класс не содержит подклассов и имеет следующие атрибуты и домены:

Название места,

Географический тип (балка / бульвар / бухта / гора / город / долина / кладбище / курган / монастырь / мост / мыс / остров / полуостров / поселок / проспект / район / регион / река / республика / село / страна / улица / хутор),

Название на других языках.

Класс C^{7H} описывает множество МЕМОРИАЛЬНЫЕ ОБЪЕКТЫ. Класс не содержит подклассов и имеет следующие атрибуты и домены:

Название объекта,

Место размещения,

Тип объекта (мемориальная доска / музейный экспонат / памятник или памятный знак / топоним),

Название на других языках.

Класс C^{8H} описывает множество ПЕРИОД. Класс не содержит подклассов, периодизация отражается в атрибутах и доменах класса:

Название периода,

Начало периода (год, месяц, день),

Конец периода (год, месяц, день).

Класс C^{9H} описывает множество НАУЧНЫЙ РЕЗУЛЬТАТ. Класс и его подклассы не содержат экземпляров, описание класса содержит только перечисление дочерних подклассов:

$C^{9H.1}$ – методологические проблемы историографии развития области науки или техники,

$C^{9H.2}$ – проблемы истории технических средств,

$C^{9H.3}$ – проблемы истории развития производства,

$C^{9H.4}$ – проблемы истории НИР и ОКР,

$C^{9H.5}$ – проблемы истории научных школ.

Класс C^{10H} описывает множество РАЗДЕЛ НАУКИ. Класс и его подклассы не содержат экземпляров, описание класса содержит только перечисление дочерних подклассов 1-го и 2-го уровня:

- $C^{10H.1}$ – теоретический базис,
- $C^{10H.1.2}$ – изучение и анализ наследия классиков,
- $C^{10H.1.3}$ – теоретические и методологические работы в области истории развития области науки и техники,
- $C^{10H.1.4}$ – изучение технической политики государств,
- $C^{10H.1.5}$ – теоретические и методологические основы преподавания истории развития области науки и техники,
- $C^{10H.2}$ – история научно-технической мысли и технических средств,
 - $C^{10H.2.1}$ – исследования по всеобщей истории развития области науки и техники,
 - $C^{10H.2.2}$ – исследования по всеобщей истории развития отдельных подобластей науки и техники,
 - $C^{10H.2.3}$ – история международного научно-технического сотрудничества в области науки и техники,
 - $C^{10H.2.4}$ – региональные исследования по истории развития направлений области науки и техники,
 - $C^{10H.2.5}$ – история отдельных технических средств, систем, приборов,
- $C^{10H.3}$ – исторические материалы о научно-технических учреждениях и производственных предприятиях,
 - $C^{10H.3.1}$ – история организации и проведения НИР и ОКР,
 - $C^{10H.3.2}$ – исследование по истории профильных научно-технических учреждений,
 - $C^{10H.3.3}$ – исследования по истории промышленности в целом,
 - $C^{10H.3.4}$ – исследования по истории отдельных организаций,
- $C^{10H.4}$ – фактологическая основа истории развития отрасли науки и техники,
 - $C^{10H.4.1}$ – исследования в области вещественных памятников науки и техники (история их создания и развития, описание памятников и экспонатов музеев),
 - $C^{10H.4.2}$ – исследования по музейному делу в области науки и техники (структура музеев, поиск и научная обработка экспонатов, реставрационная деятельность),

$C^{10H.4.3}$ – публикация отдельных документов и материалов, описания отдельных фактов истории развития области науки и техники,

$C^{10H.4.4}$ – исследование эпистолярного наследия ученых и специалистов,

$C^{10H.4.5}$ – исследование мемуарной литературы,

$C^{10H.4.6}$ – изучение и пополнение научно-биографической литературы (в т. ч. некрологи),

$C^{10H.5}$ – источниковедческие основы истории развития науки и техники,

$C^{10H.5.1}$ – источниковедческие исследования по истории развития области науки и техники,

$C^{10H.5.2}$ – изучение и создание библиографической и справочной литературы: каталогов, списков трудов, тематических указателей, указателей персоналий, биографических словарей и энциклопедий, хронологических указателей, календарей знаменательных дат и др.

Класс C^{11} описывает множество ОБЪЕКТ ИССЛЕДОВАНИЯ. Описание класса должно содержать только перечисление дочерних подклассов 1-го и 2-го уровня, классификация и иерархия которых м. б. составлена на основе тезауруса [8] или его отечественного аналога [9]. Основаниями для такого использования являются следующее.

1. Упомянутые выше тезаурусы хоть и не являются «всеохватывающими», как библиотечные и аналогичные им системы, представляют достаточно крупный кластер метазнаний, которым является научно-техническая терминология. Терминологическое однообразие в пределах такого кластера является важным аргументом в пользу использования их в качестве основы для построения порталов знаний во многих областях науки и техники.

2. Построение тезаурусов, при котором учитываются не только собственно объекты исследования, но и важные системные аспекты, такие как исследования, разработки, испытания (10-я дескрипторная область) и элементы 31-й дескрипторной области: надежность, эксплуатационные характеристики (3101), техническая эксплуатация (3103) и др.

3. Достаточно высокая временная терминологическая стабильность тезаурусов на верхних уровнях (на уровнях дескрипторных областей и групп) [4].

Для разрабатываемой онтологии актуальными являются следующие отношения и их атрибуты, специфицирующие связь между аргументами.

Отношения наследования R_T (передача атрибутов и доменов родительского класса дочернему) и включения R_p (установление отношений «часть – целое») используются в следующих классах и соответствующих им подклассах (см. рис. 1):

$$\begin{aligned}
R_{T(C^1)}, R_{P(C^1)} &\rightarrow R_{T(C^{1.i})}, R_{P(C^{1.i})}, \\
R_{T(C^{2H})}, R_{P(C^{2H})} &\rightarrow R_{T(C^{2H.i})}, R_{P(C^{2H.i})}, \\
R_{P(C^3)} &\rightarrow R_{P(C^{3.1})}, \\
R_{T(C^{9H})}, R_{P(C^{9H})} &\rightarrow R_{T(C^{9H.i})}, R_{P(C^{9H.i})}, \\
R_{T(C^{10H})}, R_{P(C^{10H})} &\rightarrow R_{T(C^{10H.i})}, R_{P(C^{10H.i})} \rightarrow R_{T(C^{10H.i.j})}, R_{P(C^{10H.i.j})}, \\
R_{T(C^{11})}, R_{P(C^{11})} &\rightarrow R_{T(C^{11.i})}, R_{P(C^{11.i.j})} \rightarrow R_{T(C^{11.i.j})}, R_{P(C^{11.i.j})}.
\end{aligned}$$

Ассоциативные отношения рассмотрим в контексте атрибутов, характерных для рассматриваемой онтологии. Из 55 возможных ассоциативных отношений между классами выделим 30 наиболее актуальных (см. рис. 1):

1. *исследователь_метод* используется для установления связи между исследователем и используемым им методом исследования

$$R_{A_1} = \{C^1 \times C^{7H}\};$$

2. *автор_публикация* используется для установления связи между исследователем, являющимся автором публикации, и самой публикацией, которая является элементом класса C^3

$$R_{A_2} = \{C^1 \times C^3\};$$

3. *исследователь_работает_в* связывает исследователя с представляемой им организацией

$$R_{A_3} = \{C^1 \times C^5\};$$

4. *исследователь_описывает_мемор_объект* связывает исследователя с описываемым им мемориальным объектом

$$R_{A_4} = \{C^1 \times C^{7H}\};$$

5. *исследователем_получен_научный_результат* связывает исследователя с полученным им научным результатом

$$R_{A_5} = \{C^1 \times C^{9H.i}\};$$

6. *автором_исследуется_раздел_науки* связывает исследователя с разделом науки

$$R_{A_6} = \{C^1 \times C^{10H.i}\} \cup \{C^1 \times C^{10H.i.j}\};$$

7. *автором_исследуется_объект* связывает исследователя с объектом исследования

$$R_{A_7} = \{C^1 \times C^{11H.i}\} \cup \{C^1 \times C^{11H.i.j}\};$$

$$R_{A_7} = \{C^1 \times C^{11H,i}\} \cup \{C^{11H,i,j}\};$$

8. *события описаны в* связывает события с источником, в котором описаны эти события

$$R_{A_8} = \{C^4 \times C^3\} \cup \{C^{3,1}\};$$

9. *издано в* связывает источник (публикацию) с издательством (организацией)

$$R_{A_9} = \{C^3 \times C^5\};$$

10. *мемориальный объект описан в* связывает мемориальный объект с источником, в котором описан этот объект

$$R_{A_{10}} = \{C^{7H} \times C^3\} \cup \{C^{3,1}\};$$

11. *период описан в источниках* связывает период и его описание в источниках

$$R_{A_{11}} = \{C^3 \times C^{8H}\};$$

12. *раздел науки описан в источниках* связывает раздел науки и его описание в источниках

$$R_{A_{12}} = \{C^3 \times C^{10H,i}\} \cup \{C^{10H,i,j}\};$$

13. *объект исследования описан в источниках* связывает объект исследования и его описание в источниках

$$R_{A_{13}} = \{C^3 \times C^{11,i}\} \cup \{C^{11,i,j}\};$$

14. *событие произошло в организации* связывает организацию с происшедшим в ней событием

$$R_{A_{14}} = \{C^{4,i} \times C^5\};$$

15. *событие произошло в* связывает событие с географическим местом, где оно произошло

$$R_{A_{15}} = \{C^{4,i} \times C^6\};$$

16. *мемориальный объект посвящен событию* связывает событие с посвященным ему мемориальным объектом

$$R_{A_{16}} = \{C^{4,i} \times C^{7H}\};$$

17. *событие произошло в период* связывает событие с периодом, в который оно произошло

$$R_{A_{17}} = \{C^{4,i} \times C^{8H}\};$$

18. *событие относится к объекту* связывает событие с объектом исследования

$$R_{A_{18}} = \{C^{4,i} \times C^{11H,i}\} \cup \{C^{11H,i,j}\};$$

19. *организация и мемориальный объект* связывает мемориальный объект с организацией, имеющей отношение к его созданию

$$R_{A_9} = \{C^5 \times C^{7H}\};$$

20. *в организации получен научный результат* связывает организацию с полученным в ней научным результатом

$$R_{A_{20}} = \{C^5 \times C^{9H.i}\};$$

21. *в организации исследуется раздел науки* связывает организацию с исследуемым в ней разделом науки

$$R_{A_{21}} = \{C^5 \times C^{10H.i}\} \cup \{C^{10H.i,j}\};$$

22. *в организации исследуется объект* связывает организацию с изучаемым в ней объектом исследования

$$R_{A_{22}} = \{C^5 \times C^{11H.i}\} \cup \{C^{11H.i,j}\};$$

23. *мемориальный объект расположен в* связывает мемориальный объект с географическим местом его расположения

$$R_{A_{23}} = \{C^6 \times C^{7H}\};$$

24. *объект исследования расположен в* связывает объект исследования с географическим местом его расположения

$$R_{A_{24}} = \{C^6 \times C^{11H.i,j}\};$$

25. *мемориальный объект создан в период* связывает мемориальный объект с периодом, в который был создан этот объект

$$R_{A_{25}} = \{C^{7H} \times C^{8H}\};$$

26. *мемориальный объект посвящен объекту* связывает мемориальный объект с объектом исследования, которому посвящен мемориальный объект

$$R_{A_{26}} = \{C^{7H} \times C^{11H.i,j}\};$$

27. *объект исследования создан в период* связывает объект исследования и период его создания

$$R_{A_{27}} = \{C^{8H} \times C^{11H.i,j}\};$$

28. *научн результат получен в разделе науки* связывает раздел науки и полученный научный результат

$$R_{A_{28}} = \{C^{9H.i} \times C^{10H.i}\} \cup \{C^{10H.i,j}\};$$

29. *научн результат получен применительно к* связывает объект исследования и полученный научный результат

$$R_{A_{29}} = \{C^{9H.i} \times C^{11H.i}\} \cup \{C^{11H.i,j}\};$$

30. *раздел науки исследует объект* связывает раздел науки и объект исследования

$$R_{A_{30}} = \{C^{10H,i} \times C^{11H,i}\} \cup \dots \cup \{C^{11H,i,j}\}.$$

Определенные выше ассоциативные отношения отражают только основные бинарные связи и могут быть транзитивно замкнуты на другие объекты онтологии.

Отношения класс – данные, которые отражают связь класса C_i^3 – ИСТОЧНИКИ и его подкласса $C_i^{3,1}$ – Цитаты с множеством элементов контента портала S – SOURCES и Cit – Citations, могут быть представлены как

$$R_{CD} = \{C_i^3 \times S_i\} \cup \dots \cup \{Cit_j\}.$$

Таким образом, предлагаемый подход является инструментом для создания портала знаний в области историографии науки и техники. Для этого необходимо использовать разработанные в настоящей статье элементы базовой онтологии с соответствующим дополнением элементами предметной онтологии, касающимся конкретной области истории науки и техники.

Для реализации портала знаний в области истории науки и техники могут быть привлечены ориентированные на решение таких задач программные продукты, например, разработанная в Институте систем информатики им. А. П. Ершова СО РАН программная оболочка для построения порталов знаний, основанных на онтологиях [10]. Этот продукт имеет интуитивно понятный интерфейс и снабжен функцией поиска.

Дальнейшие исследования автора по этому направлению будут связаны с реализацией портала знаний по истории развития радиотехнологий в Крыму.

Автор выражает благодарность академику НАН Украины Ильченко М. Е. за внимание к тематике исследований.

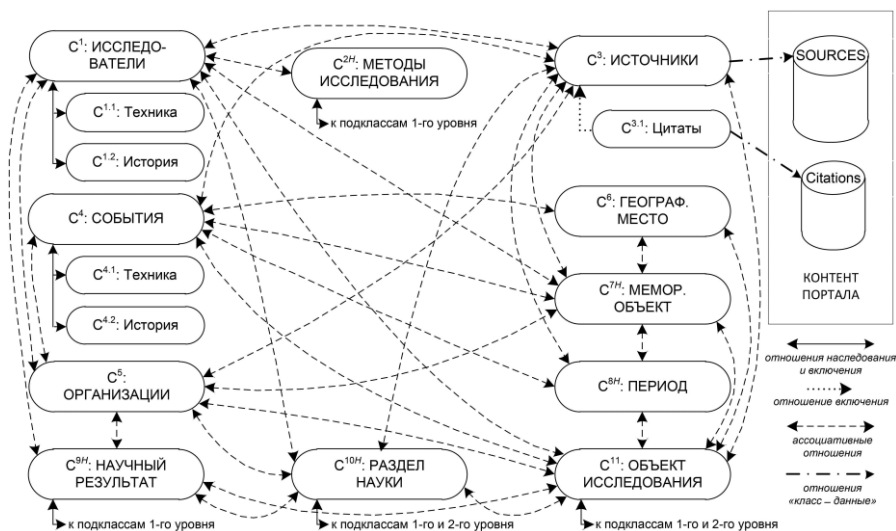


Рис. 1. Базовая онтология портала знаний по истории науки и техники

Список литературы: 1. *Щеглов С. Н.* Онтологический подход и его использование в системах представления знаний // *Известия Южного федерального университета. Технические науки.* 2009. Т. 93. № 4. С. 146–153. 2. *Загоруйко Ю. А.* Построение порталов научных знаний на основе онтологий // *Вычислительные технологии : спец. выпуск.* 2007. Т. 12. С. 169–177. 3. *Смоленский Н. И.* Теория и методология истории : учеб. пособие для студ. высш. учеб. заведений. 2-е изд., стер. М. : Издательский центр «Академия», 2008. 272 с. 4. *Ермолов П. П.* Предметные онтологии в общей и региональной историографии развития радиотехнологий // 20-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии» (КрыМиКо'2010) : материалы конф. в 2 т. Севастополь, 13–17 сент. 2010 г. Севастополь : Вебер, 2010. Т. 1. С. 73–78. 5. *Ржевцева Н. Л.* Библиометрический анализ публикаций международной конференции КрыМиКо (1991–2008 гг.) на примере Беларуси // *Інновації і менеджмент якості в діяльності бібліотек вищих навчальних закладів : матеріали V наук.-практ. конф., [22–23 жовт. 2009 р.] / Донец. нац. ун-т економіки і торгівлі ім. Михайла Туган-Барановського.* Донецьк, 2009. С. 127–135. 6. *Ермолов П. П.* Историография науки и техники как метанаука // *Вісник Національного технічного університету «Харківський політехнічний інститут».* Тематичний випуск : Історія науки і техніки. Х. : 2012. № 42. С. 23–30. 7. *Загоруйко Ю. А.* Подход к построению интеллектуальных информационных систем на основе семантических сетей // *Материалы Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2011).* Минск, 10–12 февраля 2011 г. С. 15–20. 8. *Thesaurus of Engineering and Scientific Terms. A List of Engineering and Related Scientific Terms and their Relationship for Use as a Vocabulary Reference in Indexing and Retrieving Technical Information.* Washington : Dept. of Defence, 1969. 690 p. 9. *Тезаурус научно-технических терминов / Под ред. Ю. И. Шемакина.* М. : Воениздат, 1972. 672 с. 10. *Программная оболочка для построения порталов знаний, основанных на онтологиях.* Св-во Роспатента № 2011618756 (9.11.2011) / Загоруйко Ю. А., Андреева О. А. и др.

Поступила в редколлегию 12.12.2012

Формализация структуры портала знаний по истории науки и техники / П. П. Ермолов // Вісник НТУ «ХПІ». Серія: Історія науки і техніки. – Х. : НТУ «ХПІ», 2013. – № 68 (1041). – С. 62–76. – Бібліогр.: 10 назв.

Проведено формалізований опис структури порталу знань з історії науки і техніки. Побудована базова онтологія порталу, в яку входять 11 класів: «дослідники», «методи дослідження», «джерела», «події», «організації», «географічне місце», «меморіальний об'єкт», «період», «науковий результат», «розділ науки» та «об'єкт дослідження». Останні три класи представлені як метапоняття предметної онтології порталу. На класах і підкласах визначені відносини спадкування, включення, відносини «клас — дані», виділені 30 найбільш актуальних асоціативних відносин. Для спрощення створення онтології та реалізації релевантного пошуку запропоновано поділ значних за обсягом джерел (монографій та обзорів) на більш дрібні фрагменти і організація підкласу «цитати». При проведенні формалізації використані апарат теорії множин і «двошарова» концепція історіографії науки і техніки як метанауки.

Ключові слова: институт, ученый, теория машин, теплофизика, механика, энергетика, энергетическое машиностроение, турбины.

Formal description has been developed for the structure of knowledge portal concerning history of science and technology. The basic portal ontology has been built. It consists of eleven classes: "researchers", "researching techniques", "sources", "events", "organizations", "geographical location", "memorial object", "period", "scientific result", "branch of science" and "subject under investigation". The last three classes are presented as metaconcepts of subject ontology of the portal. Dependences of inheritance and implication, as well as dependences "class — data", have been determined using classes and subclasses. Thirty most actual associative dependences have been emphasized. In order to simplify the process of ontology development and realization of the most relevant search, it is proposed to divide the large sources (monographs and reviews) into smaller fragments, and "quotations" subclass organization. During the process of formalization, set theory has been used, as well as "two-level" concept of historiography of science and technology as metascience.

Keywords: Institute scientist, theory of machines, thermal, mechanical, power engineering, energy engineering, turbine